



## بررسی تنوع ساختار کروموزومی در ژنوم گاوهای هلشتاین با استفاده از بسته نشانگری 50K

مریم نصرتی<sup>۱\*</sup>، مجتبی طهمورث پور<sup>۲</sup>، محمدرضا نصیری<sup>۳</sup>

<sup>۱</sup> دانشجوی دکتری رشته ژنتیک و اصلاح نژاد دام گروه علوم دامی، دانشکده کشاورزی، دانشگاه فردوسی مشهد.

<sup>۲</sup> استاد رشته ژنتیک و اصلاح نژاد دام گروه علوم دامی، دانشکده کشاورزی، دانشگاه فردوسی مشهد.

<sup>۳</sup> دانشیار رشته ژنتیک و اصلاح نژاد دام گروه علوم دامی، دانشکده کشاورزی، دانشگاه فردوسی مشهد

تاریخ دریافت: ۱۳۹۱/۰۷/۱۶، تاریخ پذیرش: ۱۳۹۲/۰۴/۳۰

### چکیده

تنوع ساختار کروموزومی در مکانیزم‌های بیولوژیکی از اهمیت خاصی برخوردار است. به منظور تعیین تعداد و توزیع این تنوع در ژنوم سه کروموزوم ۶، ۱۴ و ۲۰ که حضور جایگاه صفات کمی موثر بر تولید در آن تایید شده است، مورد بررسی قرار گرفت. به این منظور از ۵۸۰ راس گاو هلشتاین نمونه خون گرفته شد و پس از استخراج DNA، تعیین ژنوتیپ با استفاده از بسته نشانگری 50K گاو شرکت ایلومینا صورت گرفت. پس از تصحیح داده‌ها برای شدت سیگنال و نواحی غنی از GC، تعداد ۳۸۳ نمونه برای آنالیز نهایی باقی ماند. داده‌ها بر اساس اسمبلی UMD3.1 ژنوم گاو برای کروموزوم‌های ۶، ۱۴ و ۲۰ آنالیز شد. پس از اعمال کلیه فیلترها تعداد ۱۹۹ تنوع ساختار کروموزومی (۱۳۲ حذف و ۶۷ اضافه) با متوسط ۰/۵ برای هر فرد و میانگین طول ۱۴۷/۳ kb و میانه ۱۳۹/۴ kb شناسایی شد. نسبت حذف به اضافه برابر با ۱/۹۷ و کروموزوم‌های ۶ و ۲۰ به ترتیب دارای بیشترین و کمترین تنوع ساختار کروموزومی بودند. آنالیزهای بیوانفورماتیکی مشخص نمود که ۷۷ ژن با این نواحی تنوع ساختار کروموزومی در ارتباط هستند. نتایج این پژوهش نشان داد تنوع ساختار کروموزومی بخش قابل توجهی از این سه کروموزوم که حاوی تعداد قابل توجهی جایگاه صفات کمی هستند را پوشش می‌دهد. به دلیل اثر این تنوع بر ایجاد تغییر در ساختار و دز ژن‌ها، به نظر می‌رسد تنوع ساختار کروموزومی می‌تواند بر واریانس فنوتیپی صفات کمی موثر باشد لذا احتمالاً این تنوع قابلیت استفاده در برنامه‌های اصلاح نژادی را دارد.

**کلید واژه:** گاو هلشتاین، تنوع ساختار کروموزومی، جهش تک نوکلئوتیدی، ژنوم.

## مقدمه

فنوتیپی بیش از جهش‌های تک نوکلئوتیدی می‌باشد (Li et al, 2010). در پستانداران این تنوع ۰/۴-۵٪ کل تنوع ژنوم را شامل می‌شود (Fontanesi et al, 2010). اولین مطالعات بر روی تنوع ساختار کروموزومی در انسان و در سال ۲۰۰۶ صورت گرفت (Roden et al. 2006) از آن زمان تاکنون ۲۹۰۰۰ تنوع ساختار کروموزومی در انسان شناسایی شده است که ۹۰۰۰ مورد آن در نزدیکی یا داخل یک ژن قرار دارند (Conrad et al, 2009; Redon et al, 2006). ۱۲٪ تنوع ژنوم انسان مرتبط با تنوع در ساختار کروموزومی است و تقریباً ۴۰٪ درصد این ساختارها حداقل با یک ژن همپوشانی دارد (Li et al 2010). به علاوه، این نوع تنوع با بیماری‌های از قبیل شیذوفرنی، دیابت نوع اول و سرطان و ... مرتبط است (Kijas et al 2011; Hastings et al 2011; Mills et al 2009; et al). با این حال، مطالعه روی نحوه گسترش، توزیع و نقش تنوع ساختار کروموزومی در حیوانات اهلی در مراحل ابتدایی است. این امر می‌تواند به دلیل کمبود ابزارهای ژنومی در حیوانات اهلی باشد. اولین مطالعه برای شناسایی تنوع ساختار کروموزومی گاو توسط Liu et al. (2008) با روش aCGH<sup>3</sup> صورت گرفت. آنها توانستند ۲۵ تنوع ساختار کروموزومی را در سلول‌های جنسی شناسایی کنند (Liu et al 2008). پس از آن مطالعات روی

گونه گاو در حدود ۱۰۰۰۰ سال پیش در مناطقی از آسیا و خاور نزدیک اهلی شد (Gotherstrom et al. 2005). بدلیل اهمیت آن در تولید شیر، گوشت و فعالیت‌های کشاورزی مطالعه برای پرورش و اصلاح نژاد آن گسترش یافت. ژنوم گاو از ۲۹ جفت کروموزوم اتوزوم و یک جفت کروموزوم جنسی تشکیل شده است که در حدود ۲/۸۷ Gbp طول دارد. تاکنون ۲۲۰۰۰ ژن در آن شناسایی شده است (Bae et al 2010). بر اساس مطالعات مختلف، جایگاه بسیاری از صفات کمی<sup>۱</sup> مانند رشد، تولید شیر و ترکیب آن و ماربلینگ روی کروموزوم ۱۴، ۶ و ۲۰ شناسایی شده است (Plante et al, 2001; Miyata et al, 2007; Mehar et al, 2004; Marques et al, 2008). لذا مطالعه روی تنوع ژنومی در این کروموزوم‌ها از اهمیت خاصی برخوردار است.

انواع تنوع ژنومی شامل حذف، اضافه، جایگزینی، بازترتیبی، معکوس شدن و تنوع ساختار کروموزومی<sup>۲</sup> (CNV) است (liu 2011). تنوع ساختار کروموزومی شامل حذف، اضافه یا دو برابر شدن قطعه‌ای از DNA به طول ۱ kb تا چندین مگا جفت باز است. مطالعات نشان می‌دهد که این تنوع می‌تواند منجر به تغییر در دز ژن، توالی‌های کد کننده و تنظیم بیان ژن شود (Aboura et al 2009). لذا تاثیر آن بر تنوع

<sup>3</sup> Array-comparative genomic hybridization

<sup>1</sup> Quantitative traits loci

<sup>2</sup> Copy Number Variance

گاوه هلشتاین با استفاده از بسته نشانگری 50K گاوی روی کروموزوم‌های ۶، ۱۴ و ۲۰ مورد بررسی قرار گرفت.

#### مواد و روشها

##### نمونه‌گیری و تعیین ژنوتیپ

نمونه خون از ۵۸۰ گاوه هلشتاین دارای شجره از موسسه ANAFI<sup>۱</sup> تهیه شد. استخراج DNA از خون با استفاده از کیت Wizard<sup>®</sup> Genomic DNA Purification kit (Promega Corporation, Madison, WI) ژنتیک مرکز ژنومیک حیوانات اهلی، دانشگاه بولونیا کشور ایتالیا انجام شد. در این پژوهش بسته نشانگری 50K گاوی نوع دو شرکت ایلومینا<sup>۲</sup> (Illumina, Inc., USA) برای تعیین ژنوتیپ ۵۴۶۰۴ SNP در سراسر ژنوم استفاده شد. چپ مورد نظر بر مبنای پروتکل Infinium HD Assay شرکت ایلومینا طراحی شده است. مفهوم Infinium بر اساس هیبریداسیون مستقیم قطعات کل ژنوم به پرایمرهای ۵۰ مری متصل به آرایه دانه<sup>۳</sup> می‌باشد (and Gunderson, 2007). نمونه‌های DNA با غلظت ۴۰۰ ng-۲۰۰ با محلول ۰/۱ نرمال NaOH دناتوره شده و سپس با محلول MA2 (کیت ایلومینا) خنثی شده تا برای تکثیر آماده شود. در مرحله بعد محلول MSM<sup>۴</sup> که حاوی مستر میکس است به DNA اضافه شده و سپس پلیت در آون هیبریداسیون در

سایر حیوانات مزرعه‌ای مانند جوجه گوستی (Wang et al 2010)، گوسفند (Fontanesi et al 2011) و بز (Fontanesi et al 2010) صورت گرفت. با توسعه تکنولوژی، ساخت چپ‌های شناسایی جهش‌های تک نوکلئوتیدی ژنوم، شناسایی تنوع ساختار کروموزومی با دقت و سرعت بیشتری انجام گرفته است.

مطالعات اخیر نشان می‌دهد تنوع ساختار کروموزومی با تنوع فنوتیپی برخی صفات مرتبط است. مثلاً "تاج نخودی در طیور ناشی از تنوع ساختار کروموزومی در اینترون یک ژن SOX5 می‌باشد (Wright et al 2009). در گوسفند دو برابر شدن در ژن ASIP منجر به تنظیم رنگدانه در پوشش بدن می‌شود (Norris and Whan, 2008). رنگ سفید غالب در پوشش بدن خوک ناشی از دو برابر شدن در ژن Kit می‌باشد (Giuffra et al, 2002). تاخیر در پر درآوری جوجه‌ها به دلیل دوبله ناقص در ژن‌های PRLR و SPEF2 می‌باشد (Elferink et al, 2008). دو برابر شدن ۴/۶ kb در اینترون ۶ ژن stx17 منجر به خاکستری شدن رنگ بدن در اسب با افزایش سن می‌شود (Rosengren et al, 2008). حذف ناحیه بین ژنی بطول ۱۱/۷ kb در ژنوم بز منجر به حذف شاخ می‌شود (Pailhoux et al, 2001). با توجه به پوشش قابل توجه این تنوع در ژنوم و اثر آن بر تنوع فنوتیپی برخی صفات، اهمیت مطالعه بیشتر روی این نوع تنوع آشکارتر می‌شود. در این پژوهش تنوع ساختار کروموزومی در ۵۸۰

<sup>1</sup> National Association of Italian Holstein Breeders

<sup>2</sup> Illumina BovineSNP50 BeadChip v2 bead array<sup>3</sup>

<sup>4</sup>Multi-Sample Amplification Master Mix

دمای  $37^{\circ}\text{C}$  به مدت ۲۴-۲۰ ساعت انکوبه شد تا تکثیر صورت گیرد. این روش موجب تکثیر DNA بطور یکنواخت تا چندین برابر و بدون اریب (خطا) خواهد شد. DNA توسط سیستم آنزیمی تحت کنترل به قطعات کوچکتر بریده شد. DNA هضم شده با ۲-پروپانول ۱۰۰٪ مخلوط و با سانتیفریوژ در دمای ۴ درجه سانتیگراد رسوب داده شد و دوباره با محلول بافر هیبریداسیون سوسپانسیون شد. در مرحله بعد نمونه های DNA بر روی Beadchips قرار داده شد و به مدت ۲۴-۱۶ در دمای ۴۸ در آن هیبریداسیون نگهداری شد. در این مرحله از طریق سیستم مویرگی موجود در اطراف چیپ و تزریق محلول های هیبریداسیون، ابتدا DNA دناتوره شده و سپس هیبرید می شود. پس از هیبریداسیون، شستشو Beadchips برای حذف DNA هیبرید نشده انجام شد. نمونه ها برای بسط تک بازی و رنگ آمیزی آماده شد. به دلیل اینکه طول پرایمرهای (یک جفت) متصل به آرایه دانه تا یک نوکلئوتید قبل از محل جهش است پس از اتصال DNA تک رشته (بعنوان الگو) به پروب ۵۰ مری (بعنوان پرایمر)، بسط تک بازی در انتهای ۳' پروب انجام می شود. بدین منظور محلول TEM در اطراف Beadchips رها شده تا با استفاده از نوکلئوتیدهای نشاندار (فلورسنس) تکثیر صورت گیرد. در این محلول A و T به رنگ قرمز و C و G به رنگ سبز نشاندار شده است. پس از این مرحله DNA الگو توسط محلول فرماید ۹۵٪ بر ۱ میکرومول EDTA

حذف شد. در این حالت پرایمر بسط یافته متصل به آرایه دانه، تحت فرآیندهای مختلف انجام گرفت تا برای اسکن آماده شود. Beadchips با استفاده از iScanReader یا Bead Array reader تحت لیزر قرار گرفت تا فلورسنس تک باز انتهای پرایمرهای بسط داده شده در آرایه دانه برانگیخته شود. اگر لوکاس مورد نظر فاقد جهش (A/T یا C/G) باشد رنگ قرمز یا سبز ساطع می کند اما اگر جهش صورت گرفته باشد یعنی در مقابل A مثلاً C باشد رنگ زرد (ترکیب قرمز و سبز) ساطع می گردد. نور ساطع شده ثبت شد و داده ها بدست آمده از این تصویر با Illumina's BeadStudio v1.0 تصحیح شد. بدین معنی که گراف تشکیل شده از SNP ها برای کلیه نمونه ها کنترل شد و دسته بندی<sup>۱</sup> و نرمال سازی<sup>۲</sup> داده ها انجام گرفت. بر اساس مطالعات Diskin et al. (2008) گاهی در هنگام هیبریداسیون بدلیل نوسان در غلظت DNA نمونه های مختلف و محتوی GC متفاوت، اتصال پروبها بطور یکنواخت صورت نمی گیرد این پدیده را امواج ژنومیکی<sup>۳</sup> می نامند که بر بازخوانی ژنوتیپها اثر می گذارد. این پدیده احتمالاً<sup>۴</sup> بر صحت تعیین تنوع ساختار کروموزومی نیز موثر است. بدین منظور GenCall Score نمونه ها بررسی شد. این پارامتر که توسط BeadStudio بدست می آید یک کمیت متریک برای نشان دادن قابلیت اطمینان<sup>۴</sup> ژنوتیپهای خوانده شده است که حد آستانه آن

<sup>1</sup> clustering

<sup>2</sup> Normalization

<sup>3</sup> Genomic wave

<sup>4</sup> reliability

شدت سیگنال نرمال شده برای هر دو آلل SNP می‌باشد که در زمان حذف، کاهش و در زمان اضافه شدن، افزایش می‌یابد. نحوه محاسبه آن به این صورت است اگر  $R = Y + X$  باشد آنگاه  $LRR = \text{Log}_2(R_{\text{observed}} / R_{\text{expected}})$  است که  $R_{\text{expected}}$  از طریق درونیابی خطی دسته‌های ژنتیکی SNP نمونه‌ها به دست می‌آید. BAF نرمال شده، نسبت شدت سیگنال نسبی آلل A و B است که از طریق محاسبه  $\theta = \tan^{-1}(Y/X) / (\pi/2)$  در شروع آنالیز، فایل شدت سیگنال ورودی به فایل‌های کوچکتر شکسته شد که هر فایل حاوی اطلاعات فوق برای هر فرد به طور مجزا است. فایل فراوانی آللی B در جمعیت (PFB) بر اساس BAF هر مارکر در جمعیت محاسبه شد. محل و موقعیت SNP ها روی کروموزوم از فایل PFB اخذ می‌شود لذا در صورت تغییر در اسمبلی ژنوم تعیین محل جهش‌ها به راحتی امکان پذیر است. فایل GC مدل بر اساس محتوی GC یک مگا جفت باز از ژنوم ساخته شد. فایل HMM مدل<sup>۴</sup> را برای نرم افزار PennCNV فراهم می‌کند و مقدار شدت سیگنال مورد انتظار برای CNV های مختلف را به برنامه القاء می‌کند. با توجه به اینکه مدل HMM برای همه ژنوم‌ها تقریباً یکسان است از فایل hh550.hmm که برگرفته از پانل Illumina HumanHap550 انسانی است استفاده شد. به این ترتیب تنوع ساختار کروموزومی با استفاده از مدل زنجیره مارکوف در نرم افزار

۰/۱۵ است. ژنوتیپ‌ها با GenCall Score کمتر از ۰/۱۵ بدلیل عدم اطمینان در تعیین ژنوتیپ صحیح حذف شد. پس از حذف نمونه‌ها ناصحیح ۴۷۲ نمونه باقی ماند که پس از انتخاب آیت‌ها مورد نیاز برای محاسبه CNV، فایل شدت سیگنال در فرمت متنی که همان فایل ورودی نرم افزار PennCNV است تهیه شد. کلیه مراحل تعیین ژنوتیپ با استفاده از کیت شرکت ایلومینا و براساس پروتکل Infinium HD Assay انجام شد. اطلاعات بیشتر در وب سایت شرکت ایلومینا (<http://www.illumina.com>) در دسترس می‌باشد.

#### آنالیز با PennCNV و کنترل کیفیت

نرم افزارهای متعددی برای شناسایی تنوع ساختار کروموزومی بر اساس داده‌های SNP وجود دارد. متداول‌ترین آنها PennCNV می‌باشد که از طریق وب سایت <http://www.neurogenome.org/cnv/penncnv> قابل دسترسی می‌باشد. فایل ورودی نرم افزار PennCNV که همان فایل شدت سیگنال<sup>۱</sup> است، حاوی اطلاعات نام SNP و موقعیت آن، پارامتر-های لگاریتم نسبت شدت سیگنال<sup>۲</sup> (LRR) و نسبت شدت سیگنال نسبی دو آلل SNP<sup>۳</sup> (BAF) می‌باشد. با توجه به نحوه تعیین ژنوتیپ SNP ها، از شدت سیگنال ساطع شده آلل A و B (مقادیر X و Y) برای تهیه فایل‌های LRR و BAF استفاده می‌شود. به این ترتیب که LRR کل

<sup>1</sup> signal intensity file

<sup>2</sup>Log R Ratio

<sup>3</sup>B Allele Frequency

<sup>4</sup> hidden Markov model

در این ناحیه شناسایی نشده است (Liu *et al* 2010; Bae *et al* 2010; Fadista *et al* 2010). در مطالعه Seroussi *et al.* (2010) برای بررسی تنوع ساختار کروموزومی ژنوم گاو این ناحیه بعنوان ژن مرجع برای آزمایش PCR کمی در نظر گرفته شده است. پرایمرها با نرم افزار Primer version 3.0 برای نواحی مورد نظر طراحی شد که طولی کمتر از ۱۵۰ kb داشت (جدول ۴). شرایط واکنش به شرح ذیل بود: ۲ میکرولیتر DNA، ۱۰ میکرولیتر محلول ۲ X UltraSYBR، ۰/۴ میکرولیتر پرایمرهای رفت (۱۰ X) و ۰/۴ میکرولیتر پرایمرهای برگشت (۱۰ X) در حجم نهایی ۲۰ میکرولیتر در شرایط دمایی ۱۰ دقیقه ۹۵°C در ۴۰ سیکل PCR با مراحل ۹۵°C بمدت ۱۵ ثانیه، ۶۰°C بمدت یک دقیقه تکرار شد. پرایمرها با منحنی دمای ذوب برای بررسی حضور دایمر پرایمر کنترل شد. آنالیزهای PCR کمی با روش  $\Delta\Delta C_t$  انجام شد. متوسط  $C_t$  برای نمونه‌های هدف (سه تکرار) محاسبه شد و  $\Delta C_t$  نمونه هدف از اختلاف متوسط  $C_t$  هر ناحیه تنوع در نمونه‌های که حضور تنوع در آن از طریق نرم افزار PennCNV تایید شده است با ناحیه ژن مرجع به دست آمد.  $\Delta C_t$  نرمال از اختلاف متوسط  $C_t$  در ناحیه مورد نظر در نمونه‌ای فاقد تنوع و ناحیه ژن مرجع محاسبه شد.  $\Delta\Delta C_t$  از اختلاف بین  $\Delta C_t$  نمونه‌های هدف و  $\Delta C_t$  نمونه‌های نرمال محاسبه شد. کمیت نسبی از رابطه  $2^{-\Delta\Delta C_t}$  بدست آمد. نرم افزار PennCNV تعداد کپی‌ها را بر اساس موجود دیپلوئید (که دارای یک جفت

PennCNV محاسبه شد (Wang *et al*, 2007). پارامتر  $cn^1$  گزارش شده توسط نرم‌افزار PennCNV در کروموزوم اتوزوم برای نواحی که فاقد تکرار هستند برابر ۲ (موجود دیپلوئید) برای نواحی که حذف صورت گرفته برابر ۱ یا صفر بسته به میزان حذف و برای نواحی دو یا سه برابر شده معادل به ترتیب ۳ و ۴ است. پس از محاسبه تنوع ساختار کروموزومی فیلترهای کنترل کیفیت نرم‌افزار PennCNV اعمال شد. نمونه‌های با انحراف استاندارد LRR کمتر از ۰/۰۳، BAF برابر با حداقل ۰/۰۱ و waviness factor value بین ۰/۰۵ تا -۰/۰۵ انتخاب شدند. تصحیح برای محتوی GC با استفاده از فایل GC مدل انجام شد. پس از اعمال فیلترهای مورد نظر تعداد ۳۸۳ نمونه باقی ماند. داده‌ها بر اساس اسمبلی UMD3.1 ژنوم گاو برای بررسی تنوع ساختار کروموزومی سه کروموزوم ۶، ۱۴ و ۲۰ آنالیز شد.

### ارزیابی نواحی تنوع ساختار کروموزومی

به منظور تایید نواحی شناسایی شده، چهار ناحیه تنوع و یک ناحیه کنترل انتخاب و آزمایش Quantitative PCR در سه تکرار و پنج نمونه برای هر CNVR انجام شد. ناحیه کنترل قسمتی از ژن RPP30 روی کروموزوم ۲۶ گاوی (۱۲۸۹۳۲۷۷-۱۲۸۹۳۴۰۸) بود که در مطالعات انسانی بعنوان ژن مرجع در نظر گرفته شده است (Wang *et al* 2010) و در مطالعات مختلف بر روی ژنوم گاو تاکنون تنوع ساختار کروموزومی

<sup>1</sup> Copy Number

میان ۱۳۹/۴ کیلو باز شناسایی شد. از این تعداد ۱۳۲ مورد حذف صورت گرفت که نسبت حذف به اضافه برابر با ۱/۹۷ بود. در بین کروموزوم‌ها، کروموزوم ۶ بیشترین تنوع و طول ساختار کروموزومی را داشت ولی با توجه به طول کروموزوم‌ها، کروموزوم ۱۴ دارای بیشترین درصد تنوع ساختار کروموزومی به میزان ۵/۴٪ بود (جدول ۱). اولین مطالعه روی تنوع ساختار کروموزومی ژنوم گاو در سال ۲۰۰۸ با شناسایی ۲۵ تنوع در ۱۶ کروموزوم به روش aCGH انجام شد (liu et al 2008). در سال ۲۰۱۰ در مطالعات Bae et al. (2010) با استفاده از بسته نشانگری 50K گاوی تعداد ۸۵۵ تنوع در کل ژنوم گاو شناسایی شد. در مطالعه آنها متوسط و میان تنوع ساختار کروموزومی به ترتیب ۱۴۹/۸ و ۱۷۱/۵ بود که با نتایج این پژوهش مطابقت داشت. Wang et al. (2010) با روش aCGH تعداد ۹۶ تنوع ساختار کروموزومی در ژنوم طيور شناسایی کردند که ۱۶ مگا جفت باز و معادل ۱/۳۴٪ ژنوم طيور بود که از این ۹۶ تنوع ۴۶ جایگاه در نواحی غیر کد کننده بود. در مطالعه Fontanesi et al (2010) با روش aCGH روی ژنوم بز تعداد ۱۶۱ تنوع ساختار کروموزومی با متوسط ۱۷/۹ تنوع به ازاء هر فرد شناسایی شد. علاوه بر این نتایج این پژوهش نشان داد که تقریباً ۶۰٪ تنوع‌ها طولی کمتر از ۲۰۰ kb داشتند (جدول ۲).

کروموزوم همولوگ است) محاسبه می کند. لذا، برای این که بتوان نتایج PCR کمی را با نتایج نرم افزار PennCNV مقایسه کرد بایستی مقدار کمیت نسبی بدست آمده از آزمایش PCR کمی را در عدد ۲ ضرب کرد. از این رو تعداد کپی هر ناحیه مورد آزمون از رابطه  $2^{-\Delta\Delta Ct} * 2$  یا  $2^{1+\Delta\Delta Ct}$  محاسبه شد. برای اطلاعات بیشتر به راهنمای نرم افزار PennCNV به قسمت مربوط به نحوه محاسبه تنوع ساختار کروموزومی در موجود دیپلوئید توسط نرم افزار (De novo CNV) مراجعه نمائید

#### نتایج و بحث

شيوه‌های متعددی برای شناسایی تنوع ساختار کروموزومی وجود دارد. از آن جمله می-توان شناسایی بر اساس SNP array , Comparative Genomic Hybridization , Next-Generation array (CGH) و Sequencing را نام برد. در مقایسه با دو روش دیگر روش SNP array ارزان‌تر و پوشش وسیع-تر و متراکم‌تری در ژنوم دارد ( Hou et al, 2012). در این پژوهش از تکنیک SNP array استفاده شد و بررسی‌های ذیل انجام شد.

#### شناسایی تنوع ساختار کروموزومی

پس از اعمال کلیه فیلترها تعداد ۱۹۹ جایگاه تنوع ساختار کروموزومی با متوسط ۰/۵ جایگاه برای هر فرد و میانگین طول ۱۴۷/۳ و

جدول ۱ - توزیع تنوع ساختار کروموزومی (CNV) در کروموزوم‌های ۶، ۱۴ و ۲۰ با نرم افزار PennCNV

**Table 1-Copy number variation on BTA6, BTA14, BTA20 with PennCNV.**

| درصد<br>percentage | طول<br>کروموزوم<br>Chromosome<br>length | طول تنوع<br>CNV length | اضافه<br>gain | حذف<br>loss | تعداد کل<br>تنوع<br>Total CNV | کروموزوم<br>chr |
|--------------------|---|------------------------|---------------|-------------|-------------------------------|-----------------|
| 4.26               | 119458736                               | 5085829                | 35            | 70          | 105                           | BTA6            |
| 5.4                | 84648390                                | 4569117                | 20            | 37          | 57                            | BAT14           |
| 3.85               | 72042655                                | 2770349                | 12            | 25          | 37                            | BAT20           |
| --                 | 276149781                               | 12425295               | 67            | 132         | 199                           | جمع             |

اندازه کروموزوم از سایت [www.ensembl.org](http://www.ensembl.org) بدست آمد.

جدول ۲- توزیع نواحی تنوع ساختار کروموزومی (CNVR) در کروموزوم‌های ۶، ۱۴ و ۲۰.

**Table 2- Copy number variation region (CNVR) distribution on BTA6, BTA14 and BTA20.**

| ۴۰۰ kb □ | -۴۰۰ Kb | Kb   | -۲۰۰ Kb | kb < | کل طول<br>نواحی<br>CNVR<br>length | تعداد نواحی تنوع<br>CNVR<br>number | کروموزوم<br>chromosome |
|----------|---------|------|---------|------|-----------------------------------|------------------------------------|------------------------|
|          | ۳۰۰     | -۳۰۰ | ۱۰۰     | ۱۰۰  |                                   |                                    |                        |
| 17.1     | 6.7     | 13.8 | 34.9    | 27.5 | 5.085                             | 17                                 | BTA6                   |
| 21.8     | 3.5     | 24.1 | 24.6    | 26   | 4.569                             | 12                                 | BTA14                  |
| 16.5     | 13.5    | 8    | 31      | 31   | 2.770                             | 12                                 | BTA20                  |
| --       | --      | --   | --      | --   | 12.324                            | 41                                 | جمع                    |

اعداد بر اساس درصد می‌باشد.

۱۴۶kb بود (جدول ۲). این میزان ۰/۶٪ توالی کروموزوم‌های اتوزومی و ۰/۵٪ کل ژنوم گاو را در بر می‌گیرد. بیشترین تعداد و طول نواحی تنوع ساختار کروموزومی روی کروموزوم شش مشاهده گردید. Liu et al. (2010) ۲۰۰ ناحیه تنوع ساختار کروموزومی با پوشش ۱/۰۷ درصد کل ژنوم در نژادهای مختلف گاو شناسایی کردند

### نواحی تنوع ساختار کروموزومی

نواحی تنوع ساختار کروموزومی با استفاده از یکی کردن نواحی از تنوع که با هم همپوشانی داشتند ساخته شد (Redon et al 2006). به طور کلی تعداد ۴۱ ناحیه تنوع ساختار کروموزومی در این مطالعه شناسایی شد که شامل ۱۲/۳ مگا جفت باز با میانگین و میانه به ترتیب ۳۰۳ kb و



## ارزیابی نواحی تنوع ساختار کروموزومی

به طور کلی طول ناحیه تنوع ساختار کروموزومی ژنوم گاو در مطالعات مختلف متفاوت است و از ۳۲ kb تا ۵/۶ Mb متغیر است. مقدار گزارش شده در این مطالعه در داخل این دامنه قرار دارد (۱ Mb - ۷۴ kb). *Hou et al.* (2011) دلیل این تنوع را در استفاده از روش‌های مختلف شناسایی تنوع ساختار کروموزومی یعنی aCGH و بسته نشانگری ۵۰ K گاوی اذعان کردند. بسته نشانگری ۵۰ K با تراکم بالای SNP برای مطالعات ارتباطی در ژنوم گاو طراحی شده است، لذا توزیع SNPها در همه قسمت‌های ژنوم یکنواخت نیست. شناسایی این تنوع با نرم افزار PennCNV نسبت به تراکم پایین SNPها و توزیع غیریکنواخت آن بخصوص در نواحی حاوی تنوع ساختار کروموزومی حساس است. لذا در اغلب موارد فقط تنوع‌های بلند شناسایی می‌شود. علاوه بر این بسته نشانگری ۵۰ K برخلاف aCGH فاقد پروب اختصاصی برای نواحی تنوع ساختار کروموزومی است. از این رو نتایج پژوهش‌های اخیر با مطالعات قبلی (با روش aCGH) متفاوت است و تنوع زیادی بین آنها وجود دارد. به منظور تایید تنوع ساختار کروموزومی از بین ۴۱ CNVR شناخته شده برای چهار ناحیه و یک کنترل، آزمایش PCR کمی انجام شد (جدول ۴). برای هر ناحیه از نمونه‌ای که تنوع ساختار کروموزومی توسط نرم افزار PennCNV در آن تایید شده بود به عنوان نمونه هدف و در نمونه‌هایی که حضور تنوع تایید نشده

که با ۴۰۰ ژن مختلف همپوشانی داشت. *Hou et al.* (2011) در بررسی تنوع ساختار کروموزومی گاو، تعداد ۶۸۲ ناحیه تنوع به طول کلی ۱۳۹/۸ مگاجفت باز (۴/۶٪ ژنوم) شناسایی کردند. در مطالعه *Hou et al.* (2012) تعداد ۳۳۴۶ ناحیه تنوع ساختار کروموزومی با طول کلی ۱۴۲/۷ مگاجفت باز در اسمبلی Btau\_4.0 ژنوم گاو و تعداد ۳۴۳۸ ناحیه تنوع ساختار کروموزومی با طول کلی ۱۴۶/۹ مگاجفت باز در اسمبلی UMD3.1 ژنوم گاو شناسایی کردند که به طور کلی ۴/۷٪ کل ژنوم را شامل شد. در مطالعه‌ای دیگر روی ژنوم گاو ۳۶۸ ناحیه تنوع ساختار کروموزومی شناسایی شد (*Bae et al* 2010). در پژوهش *Fontanesi et al* (2010) روی ژنوم گوسفند، با روش aCGH، ۱۳۵ ناحیه تنوع ساختار کروموزومی با طول ۱۰/۸ مگاجفت باز شناسایی شد. میزان تنوع شناسایی شده و به طبع طول آن در پژوهش‌های مختلف متفاوت است. در واقع تعداد تنوع بستگی به عوامل متعددی از قبیل روش مورد استفاده، تعداد افراد، تعداد کروموزوم-های مورد بررسی و تراکم چیپ مورد استفاده دارد. لذا با توجه به این که در این پژوهش فقط سه کروموزوم مورد بررسی قرار گرفته است تعداد کل تنوع‌ها و درصد پوشش ژنوم با سایر مطالعات متفاوت است، اما میانگین و میانه تنوع با سایر مطالعات مطابقت دارد.

کروموزومی شناسایی شد که اغلب آنها با صفات بیولوژیکی مهمی از قبیل تولید شیر، ایمنی زایی و تولید مثل در ارتباط بودند. *Fadista et al.* (2010) تعداد ۳۰۴ ناحیه تنوع ساختار کروموزومی در ژنوم گاو شناسایی کردند که ۳۰٪ آنها با ژن‌های پاسخ به محیط همپوشانی داشتند. برخی از ژن‌ها شناسایی شده در این پژوهش در مطالعات مختلف به عنوان ژن کاندیدا در تولید و رشد شناسایی شده‌اند. به عنوان مثال ژن تیروگولوبولین (TG) روی کروموزوم ۱۴ قرار دارد. مطالعات نشان می‌دهد برخی ژنوتیپ‌های آن منجر به افزایش سطح ماربلینگ عضله در گاو-های گوشتی می‌شود (Anton *et al.*, 2008; Fortes *et al.*, 2009). ژن دی‌آسیل گلیسرول آسیل‌ترانسفراز ۱ (DGAT1) که بر روی کروموزوم ۱۴ قرار دارد ارتباط معنی‌داری با ترکیب شیر و کیفیت آن دارد (Strucken *et al.*, 2010; Marques *et al.* 2011; Bouwman *et al.*, 2011). با توجه به همپوشانی این نوع تنوع با نواحی ژنی و طول قابل ملاحظه آن، در صورت واقع شدن در نواحی کد کننده یا نواحی تنظیمی به راحتی سطح بیان ژن را تحت تاثیر قرار می‌دهد. شناسایی این تنوع در ژنوم حیوانات اهلی در سال ۲۰۰۸ از ژنوم گاو شروع شده است و بررسی ارتباط آن با صفات تولیدی در حیوانات اهلی به تازگی در حال شکل‌گیری است. اما به نظر می‌رسد این نوع تنوع می‌تواند افق جدیدی را در مطالعات اصلاح‌نژادی باز کند.

بود به‌عنوان نرمال استفاده شد. حضور تنوع ساختار کروموزومی در CNVR7 روی کروموزوم ۲۰ و CNVR3 روی کروموزوم ۱۴ تایید شد ( $p < 0.05$ ). نتایج واکنش PCR کمی این پژوهش در ۵۰٪ نواحی، تنوع ساختار کروموزومی را تایید کرد (جدول ۵). در مطالعه *Fadista et al.* (2008) نیز بر روی خوک ۵۰٪ نواحی تایید شد. این مقدار در مطالعه *Hou et al.* (2011) در گاو هلشتاین ۶۰٪ بود که بیشتر از نتیجه این پژوهش می‌باشد. همچنین مقدار تایید شده در این پژوهش از مقدار گزارش شده توسط *Fadista et al.* (2011) (۶۶٪) و *Hou et al.* (2012) (۷۷٪) در گاو هلشتاین کمتر بود. این کاهش می‌تواند به دلیل تعداد کم نواحی انتخاب شده یا به دلیل خطا در برآورد به دلیل پوشش نه چندان متراکم و غیر یکنواخت SNP‌ها چپ مورد استفاده نیز باشد که باعث می‌شود طول تنوع‌ها بیش‌تر از مقدار واقعی برآورد شود.

### ارزیابی بیوانفورماتیکی نواحی تنوع ساختار کروموزومی

بررسی بیوانفورماتیکی نشان داد ۷۷ ژن در نواحی تنوع ساختار کروموزومی واقع شده است (جدول ۳). برخی از این تنوع‌ها در نواحی اگزون ژن قرار داشت. بسیاری از مطالعات حضور تنوع ساختار کروموزومی را در نواحی ژنی تایید می‌کند (*Kijas et al.*, 2011; *liu et al.*, 2008; *liu et al.*, 2010; *Fadista et al.*, 2010). در مطالعه *liu et al.* (2010) ۴۰۰ ژن در نواحی تنوع ساختار

جدول ۳- همپوشانی نواحی ساختار کروموزومی با نواحی ژنی مرجع.

Table3- Coverage of reference gene with CNVR.

| تعداد ژن<br>Gene<br>number | ژن مرجع<br>Reference gene   | طول (bp)<br>Length | ناحیه تنوع ساختار<br>کروموزومی<br>CNVR | کروموزوم<br>chromosome |
|----------------------------|---|--------------------|--|------------------------|
| 1                          | <i>PRDM5</i>  | 139384             | 4409404- 4548788                       | BTA6                   |
| 1                          | <i>UGT8</i>   | 115083             | 12119048-12234131                      |                        |
| 1                          | <i>PPA2</i>   | 74398              | 20993424-21067822                      |                        |
| 3                          | <i>ATP8A1</i> , <i>APLT</i> , <i>SHISA3</i>   | 144298             | 62788712-62933010                      |                        |
| 1                          | <i>C6H4orf22</i>  | 100774             | 97374303-97475077                      |                        |
| 3                          | <i>HGFAC</i> , <i>LRPAP1</i> , <i>NSG1</i>  | 304130             | 106981782-107285912                    |                        |
| 4                          | <i>TNIP2</i> , <i>SH3BP2</i> , <i>ADD1</i> , <i>NOP14</i> , <i>WHSC2</i><br><i>PIGG</i> , <i>FGFRL1</i> , <i>HAUS3</i> , <i>MXD4</i> , <i>RNF4</i><br><i>MFS10</i> , <i>PDE6B</i>   | 2273588            | 107678393-109951981                    |                        |
| 37                         | <i>LRRC14</i> , <i>LRRC24</i> , <i>ARHGAP39</i> , <i>COMMD5</i><br><i>PPP1R16A</i> , <i>GPT</i> , <i>MFS3</i> , <i>RECQL4</i><br><i>VPS28</i> , <i>TONSL</i> , <i>CYHR1</i> , <i>KIFC2</i> , <i>FOXH1</i><br><i>GPR172A</i> , <i>ADCK5</i> , <i>CPSF1</i> , <i>SLC39A4</i><br><i>MAF1</i> , <i>HEATR7A</i> , <i>HSF1</i> , <i>DGATI</i> , <i>SCRT1</i><br><i>EXOSC4</i> , <i>GPAA1</i> , <i>CYCI</i> , <i>SHARPIN</i><br><i>PUF60</i> , <i>NRBP2</i> , <i>GRINA</i> , <i>OPLAH</i><br><i>PYCR1</i> , <i>TSTA3</i> , <i>CCDC166</i> , <i>MAPK15</i><br><i>GSDMD</i> , <i>NAPRT1</i> , <i>EEF1D</i> | 977500             | 1514056-2491556                        | BAT14                  |
| 2                          | <i>TSNARE1</i> , <i>BAIL</i> , <i>ARC</i>   | 461705             | 2880765-3342470                        |                        |
| 3                          | <i>TRAPPC9</i> , <i>CHRA1</i> , <i>EIF2C2</i>   | 390054             | 4103850-4493904                        |                        |
| 1                          | <i>TG</i>   | 84606              | 9215622-9300228                        |                        |
| 2                          | <i>UBE2W</i> , <i>STAU2</i>   | 86362              | 39152496-39238858                      |                        |
| 8                          | <i>DSCC1</i> , <i>TAF2</i> , <i>ENPP2</i> , <i>IMPA1</i> , <i>CHMP4C</i><br><i>MTBP</i> , <i>MRPL13</i> , <i>DEPTOR</i>   | 1543819            | 83072371-84616190                      |                        |
| 1                          | <i>CDH12</i>  | 146990             | 51422910-51569900                      | BAT20                  |
| 9                          | <i>CLPTMIL</i> , <i>LPCAT1</i> , <i>NDUFS6</i> , <i>IRX4</i><br><i>SLC9A3</i> , <i>TPPP</i> , <i>BRD9</i> , <i>SLC6A18</i> , <i>TERT</i>  | 1090658            | 70644666-71735324                      |                        |
| 77                         |   |                    |  | جمع                    |

جدول ۴- توالی پرایمرها برای ارزیابی با qPCR.

Table4- the primer sequence of qPCR.

| طول قطعه<br>length | پرایمر برگشت<br>Revers primer | پرایمر رفت<br>Forward primer   | تعداد<br>کپی<br>(cn) | طول<br>ناحیه<br>CNVR | ناحیه<br>خاتمه<br>CNVR<br>End<br>point | ناحیه شروع<br>CNVR<br>Start point | کروموزوم<br>Chromo<br>som | ناحیه تنوع<br>CNVR |
|--------------------|-------------------------------|--------------------------------|----------------------|----------------------|--|-----------------------------------|---------------------------|--------------------|
| 145                | CCTTTCCTGC<br>CAGGGGTTC       | GCACACCAGG<br>CCGTGATG         | Cn=1                 | 30413<br>0           | 1072859<br>12                          | 106981782                         | 6                         | CNVR1<br>6         |
| 146                | CCTTTCCTGC<br>CAGGGGTTC       | GCACACCAGG<br>CCGTGATG         | Cn=1                 | 14699<br>0           | 5156990<br>0                           | 51422910                          | 20                        | CNVR7              |
| 133                | GAGATCGGTC<br>CTGAGCCAGC      | TGTGCGAAAC<br>TGAATTTCT<br>GC  | Cn=1                 | 10906<br>58          | 7173532<br>4                           | 70644666                          | 20                        | CNVR1<br>2         |
| 127                | GTTGTTGAGT<br>CCCGCAGAG<br>G  | CCTGACTCCT<br>TGGGACCCG        | Cn=3                 | 39005<br>4           | 4493904                                | 4103850                           | 14                        | CNVR3              |
| 131                | TGGGACCAG<br>GTTCCATGAT<br>C  | TGCTTCCATT<br>GTTTCCTGAT<br>GA | -                    | -                    | -                                      | -                                 | 26                        | Control(<br>RPP30) |

Table-5 qPCR results for CNVR in 3 chromosomes.

| تایید<br>confirm (p<0.05) | cn<br>نرم افزار<br>PennCNV | تعداد کپی<br>ها<br>(2*RQ) | RQ (2 <sup>-<br/>ΔΔCt</sup> )<br>relative<br>quantity | ΔΔC <sub>t</sub> | متوسط<br>ΔC <sub>t</sub><br>هدف | ناحیه تنوع<br>ساختار<br>کروموزومی<br>CNVR | شماره<br>number |
|---------------------------|----------------------------|---------------------------|---|------------------|---------------------------------|---|-----------------|
| خیر                       | Cn=1                       | 2.02                      | 1.01  | -0.02            | -1.22                           | CNVR16                                    | ۱               |
| بلی                       | Cn=1                       | 1.08                      | 0.54  | 0.88             | -0.32                           | CNVR7                                     | ۲               |
| خیر                       | Cn=1                       | 2.14                      | 1.09  | -0.13            | -1.33                           | CNVR12                                    | ۳               |
| بلی                       | Cn=3                       | 2.99                      | 1.49  | -0.58            | -1.78                           | CNVR3                                     | ۴               |

### نتیجه گیری

است، مطالعات ارتباطی<sup>۱</sup> انجام شود. مطالعات ارتباطی کل ژنوم به منظور شناسایی نواحی ژنومی موثر بر بیماریها (انسان) و صفات تولیدی (دام) با پانل‌های متراکم SNP انجام می‌شود. به این ترتیب که ارتباط تمامی SNP‌های ژنوم با صفت تولیدی مورد مطالعه بررسی می‌گیرد و سپس SNP‌های که ارتباط آن با صفت مورد بررسی معنی دار شده است بعنوان جهش‌های موثر بر صفت در برنامه‌های اصلاح نژادی مورد توجه قرار می‌گیرد. این مطالعات به شیوه‌های متعددی انجام می‌گیرد. قدیمی‌ترین روش آن استفاده مستقیم از خود SNP‌ها در آنالیزها بود پس از آن به دلیل تراکم بالای SNP‌های ژنوم و وقت‌گیر بودن بررسی تک تک آنها، استفاده از قطعات بزرگتر که حاوی چندین SNP و به طبع اطلاعات بیشتری هستند رایج شد. امروزه استفاده از تنوع ساختار کروموزومی نیز در این نوع مطالعات جایگاهی خاصی دارد. با توجه به این که تنوع-های ساختار کروموزومی حداقل یک کیلو باز

در این پژوهش ۱۹۹ تنوع ساختار کروموزومی و ۴۱ ناحیه با طول کلی ۱۲ مگا جفت باز در سه کروموزوم شناسایی شد (۰/۵٪). نتایج به دست آمده نشان داد، تنوع ساختار کروموزومی بخش قابل توجهی از ژنوم گاو را شامل می‌شود. با توجه به پوشش وسیع و تقریباً<sup>۱</sup> یکنواخت این تنوع در سراسر ژنوم و در بر گرفتن بخشی از ژن‌ها توسط آن، می‌توان از این تنوع در مطالعات ارتباطی استفاده کرد. اما قدم اول شناسایی محل درست و تعداد دقیق این تنوع است. استفاده از تعداد دام بیشتر و چپ‌های متراکم‌تر در این مطالعات می‌تواند تعداد تنوع ساختار کروموزوم و دقت ارزیابی را افزایش دهد. بر این اساس پیشنهاد می‌شود برای بررسی بیشتر تاثیر این تنوع بر صفات تولیدی اولاً<sup>۱</sup> از پانل‌های متراکم‌تر SNP با تعداد دام بیشتر استفاده شود، ثانیاً<sup>۱</sup> برای تایید و شناسایی نواحی از تنوع که مستقیماً<sup>۱</sup> بر تنوع فنوتیپی صفات تولیدی موثر

<sup>1</sup> Association study

### تشکر و قدردانی

نمونه‌ها و اطلاعات مورد استفاده در این پژوهش از مرکز ANAFI تهیه شد. به این وسیله مولفان از مسئولین این مرکز تشکر و قدردانی می‌نمایند.

طول دارند و می‌توانند بطور همزمان حاوی چندین TagSNP باشند لذا حامل اطلاعات بیشتری هستند و علاوه بر راحتی آنالیزها، قدرت شناسایی نواحی مرتبط با صفات کمی نیز در آنها بیشتر است، از این رو، این امکان وجود دارد که تنوع ساختار کروموزومی در این مطالعات جایگزین TagSNP‌ها شود.

### منابع

- Aboura A, Dupas C, Tachdjian G, Portnoi M, Bourcigaux N, Dewailly D, Frydman R, Fauser B, Ronci-Chaix N, Donadille B, Bouchard P, Christin-Maitre S (2009). Array Comparative Genomic Hybridization Profiling Analysis Reveals Deoxyribonucleic Acid Copy Number Variations Associated with Premature Ovarian Failure. *Journal of Clinical Endocrinology and Metabolism* 94: 4540-4546.
- Anton I, Kovacs K, Fesus L, Varhegyi J, Lehel L, Hajda Z, Polgar J P, Szabo F, Zsolnai A (2008). Effect of DGAT1 and TG gene polymorphisms on intramuscular fat and on milk production traits in different cattle breeds in Hungary. *Acta Veterinaria Hungarica Journal* 56: 181-186
- Bae J S, Cheong H S, Kim L H, NamGung S, Park T J, Chun J Y *et al.* (2010). Identification of copy number variations and common deletion polymorphisms in cattle. *BMC Genomics* 11: 232.
- Bouwman A C, Bovenhuis H, Visker M H and Arendonk J M. (2011). Genome-wide association of milk fatty acids in Dutch dairy cattle. *BMC Genetics* 12: 1471-2156
- Conrad D F, Pinto D, Redon R, Feuk L, Gokcumen O, Zhang Y, *et al* (2009). Origins and functional impact of copy number variation in the human genome. *Nature* 464: 704-712.
- Diskin, S J, Li M, Hou J, Yang S, Glessner J, Hakonarson M, Bucan M, Maris J M, and Wang K (2008). Adjustment of genomic waves in signal intensities from whole-genome SNP genotyping platforms. *Nucleic Acids Research*. 36: e126
- Elferink M G, Vallée A A, Jungerius A P, Crooijmans R P, and Groenen M A (2008). Partial duplication of the PRLR and SPEF2 genes at the late feathering locus in chicken. *BMC Genomics* 9: 391.
- Fadista J, Thomsen B, Holm L and Bendixen C (2010). Copy number variation in the bovine genome. *BMC Genomics* 11:284
- Fadista J, Nygaard M, Holm L E, B Thomsen and C Bendixen (2008). A snapshot of CNVs in the pig genome. *PLoS One* 3: e3916.
- Feuk L, Carson A R, and Scherer S W (2006). Structural variation in the human genome. *Nature Reviews Genetics* 7: 85-97
- Fontanesi L, Beretti F, Martelli P L, Colombo M, Dall'Olio S, Occidente M, Portolano B, Casadio R, Matassino D, and Russo V (2011). A first comparative map of copy number variations in the sheep genome. *Genomics* 97: 158-165

- Fontanesi L, Martelli P L , Beretti F, Riggio V, Dall'Olio S , Colombo M, Casadio R, Russo V, and Portolano B (2010). An initial comparative map of copy number variations in the goat (*Capra hircus*) genome. *BMC genomics* 11: 639.
- Fortes M S, Curi R A, Artur L, Chardulo L and Silveira A C (2009). Bovine gene polymorphisms related to fat deposition and meat tenderness. *Genetics and Molecular Biology* 32: 75-82.
- Giuffra E, Tornsten A, Marklund S, Bongcam-Rudloff E, Chardon P, Kijas J M., Anderson SI, Archibald A L., and Andersson L (2002). A large duplication associated with dominant white color in pigs originated by homologous recombination between LINE elements flanking KIT. *Mammalian Genome* 1310: 569–577.
- Gotherstrom A, Anderung C, Hellborg L, Elburg R, Smith C, Bradley D G and Ellegren H (2005). Cattle domestication in the Near East was followed by hybridization with aurochs bulls in Europe. *Proceedings of the Royal Society B* 272, 2345–2350
- Hastings P J, Lupski J R, Rosenberg S M and Ira G (2009). Mechanisms of change in gene copy number. *Nature Review Genetic* 10: 551–564.
- Hou Y, Bickhart D M, Hvinden M L, Li C, Song J, Boichard D A, Fritz S, Eggen A, DeNise S, Wiggans G R, Sonstegard T S, Van Tassell C P, and Liu G E (2012). Fine mapping of copy number variations on two cattle genome assemblies using high density SNP array. *BMC Genomics* 13: 1471-2164
- Kijas J W, Barendse W, Barris W, Harrison B, McCulloch R, McWilliam S, and Whan V (2011). Analysis of copy number variants in the cattle genome. *Gene* 482: 73–77.
- Li X, Tan L, Liu X, Lei S, Yang T, Chen X, Zhang F, Fang Y, Guo Y, Zhang L, Yan H, Pan F, Zhang Z, Peng Y, Zhou Q, He L, Zhu X, Cheng J, Zhang L, and Liu Y (2010). A genome wide association study between copy number variation (CNV) and human height in Chinese population. *Journal of Genetic & Genomics* 37: 779-785
- Liu G E, Van Tassel C P, Sonstegard T S, Li R W, Alexander L J, Keele J W, Matukumalli L K, Smith T P, and Gasbarre L C (2008). Detection of germ line and somatic copy number variations in cattle. *Journal of Developmental Biology* 132: 231-237.
- Liu G E, Hou Y, Zhu B, Cardone M F, Jiang L, *et al* (2010). Analysis of copy number variations among diverse cattle breeds. *Genome Research journal* 20: 693-703
- Liu Z J (2011). *Next Generation Sequencing and Whole Genome Selection in Aquaculture*. 1<sup>th</sup> Edition pp.87-90
- Marques E, Grant J R, Wang Z, Kolbehdari D, Stothard P, Plastow G and Moore S S (2011). Identification of candidate markers on bovine chromosome 14 (BTA14) under milk production trait quantitative trait loci in Holstein. *Journal of Animal Breeding and Genetic* 128: 305–313
- Marques E, Schnabel R D, Stothard P, Kolbehdari D, Wang Z, Taylor J F and Moore S S (2008). High density linkage disequilibrium maps of chromosome 14 in Holstein and Angus cattle. *BMC Genetics* 9: 45
- Mehar S K, Peter C T, Imke T, and Herman W R (2004). Quantitative trait loci mapping in dairy cattle: review and meta-analysis. *Genetics. Selection. Evolution*. 36, 163–190
- Mills R E, Walter K, Stewart C, Handsaker R E. *et al* (2011). Mapping copy number variation by population-scale genome sequencing. *Nature* 470: 59-65
- Miyata M, Gasparin G, Coutinho L L, Martinez M L, Machado M A, Vinicius M, Silva G B, Campos A L, Sonstegard T S, do Rosário M F, and de Almeida Regitano L C (2007). Quantitative trait loci (QTL) mapping for growth traits on bovine chromosome 14. *Genetics and Molecular Biology journal* 30: 364-369.
- Norris B J, and Whan V A (2008). a gene duplication affecting expression of the ovine ASIP gene is responsible for white and black sheep. *Genome Research journal*. 188: 1282–1293.

- Pailhoux E, Vigier B, Chaffaux S, Servel N, Taourit S, *et al* (2001). A 11.7-kb deletion triggers intersexuality and polledness in goats. *Nature Genetics*. 29: 453–458.
- Plante Y, Gibson J P, Nadesalingam J, Mehrabani-Yeganeh H, Lefebvre S, Vandervoort G, and Jansen G B (2001). Detection of Quantitative Trait Loci Affecting Milk Production Traits on 10 Chromosomes in Holstein Cattle. *Journal of Dairy Science* 84: 1516–1524
- Redon R, Ishikawa S, Fitch K R, Feuk L, Perry G H, Andrews T D, Fiegler H, Shapero M H, Carson A R, Chen W, *et al* (2006) Global variation in copy number in the human genome. *Nature* 444: 444-454.
- Rosengren P G, Golovko A, Sundstrom E, Curik I, Lennartsson J, *et al* (2008). A cis-acting regulatory mutation causes premature hair greying and susceptibility to melanoma in the horse *Nature Genetics* 40: 1004–1009.
- Seroussi, E, Glick G, Shirak A, Yakobson E, Weller J, Ezra E, Zeron Y (2010). Analysis of copy loss and gain variations in Holstein cattle autosomes using BeadChip SNPs. *BMC Genomics*. 11: 673
- Strucken E M, Rahmatalla S, Koning D D and Brockmann G A (2010). Haplotype analysis and linkage disequilibrium for DGAT1. *Archiv Tierzucht*. 53: 247-255
- Wang K, Li M, Hadley D, Liu R, Glessner J, Grant S F, Hakonarson H, and Bucan M (2007). PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Research journal* 17:1665-1674.
- Wang X, Nahashon S, Feaster T K, Bohannon-Stewart A, and Adefope N (2010). An initial map of chromosomal segmental copy number variations in the chicken. *BMC Genomics* 11: 351.
- Wright D, Boije H, Meadows J R, Bed'hom B, Gourichon D, Vieaud A, Tixier-Boichard M, Rubin C J, Inslan F, Hallbook F, and Andersson L (2009). Copy number variation in intron 1 of SOX5 causes the Pea-comb phenotype in chickens. *PLoS Genetics* 5:e1000512.

## Detection of Copy number variation in Holstein cattle genome by using 50K

Nosrati M.<sup>\*1</sup>, Tahmorespur M.<sup>2</sup>, Nassiry M.R.<sup>3</sup>

<sup>1</sup> PhD Student, Department of animal science, Faculty of Agriculture, Ferdowsi University of Mashhad, Iran.

<sup>2</sup> Professor, Department of animal science, Faculty of Agriculture, Ferdowsi University of Mashhad, Iran.

<sup>3</sup> Associate Professor, Department of animal science, Faculty of Agriculture, Ferdowsi University of Mashhad, Iran

### Abstract

Copy number variation (CNV) is important on biological mechanism. For CNV detection and distribution, three bovine autosomal chromosomes, BTA6, BTA14 and BTA6 that improved for quantitative trait loci (QTL) previously, were investigated. Blood sample we obtained from 580 animals and after DNA extraction the samples were genotyped by Illumina BovineSNP50v2 BeadChip. Three hundred eighty three samples were remained for further analysis, after correction for signal intensity and GC content. Data were analyzed based on UMD3.1 bovine genome assembly for BTA6, BTA14 and BTA20. After filtration 199 CNV were detected (132 losses and 67 gains) with 0.5 CNV per animal and mean and medium of 147.3 kb and 139.4 kb respectively. A proportion of loss to gain was 1.97 fold. The Max and Min Number of CNV detected on BTA6 and BTA20 respectively. The bioinformatics analysis showed CNV region's coverage some part of 77 reference gene in bovine genome. These results showed that CNV coverage remarkable part of the three chromosomes that contain several QTL. Because of the influence of CNV on gene dosage and gene structure, it can cause phenotypic variation in quantitative traits, so probably it can be useful on livestock breeding programs.

Keyword: Holstein cattle, CNV, SNP, Genome

\* Corresponding Author: Nosrati M.

Tel: 09153860817

Email: mehrameh58@yahoo.com